Business Gateway

Surface Coal Mining Industry Data Standards Assessment

March 3, 2006

Consulting.
Systems Integration.
Outsourcing.
Infrastructure.
Server Technology.

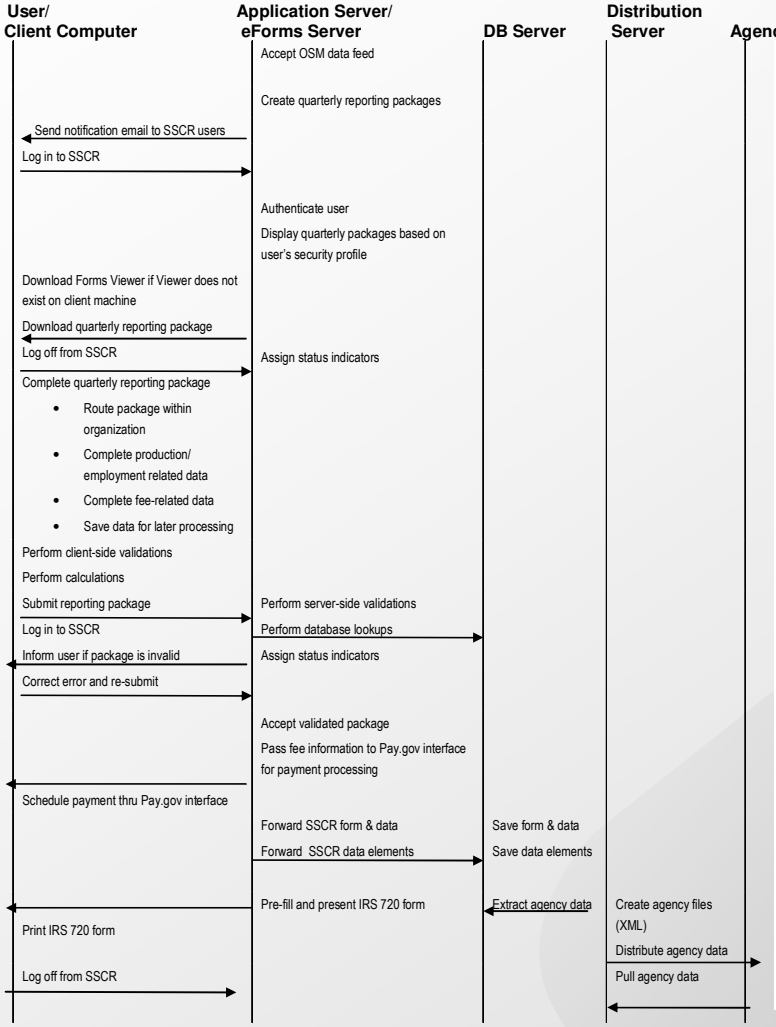
UNISYS

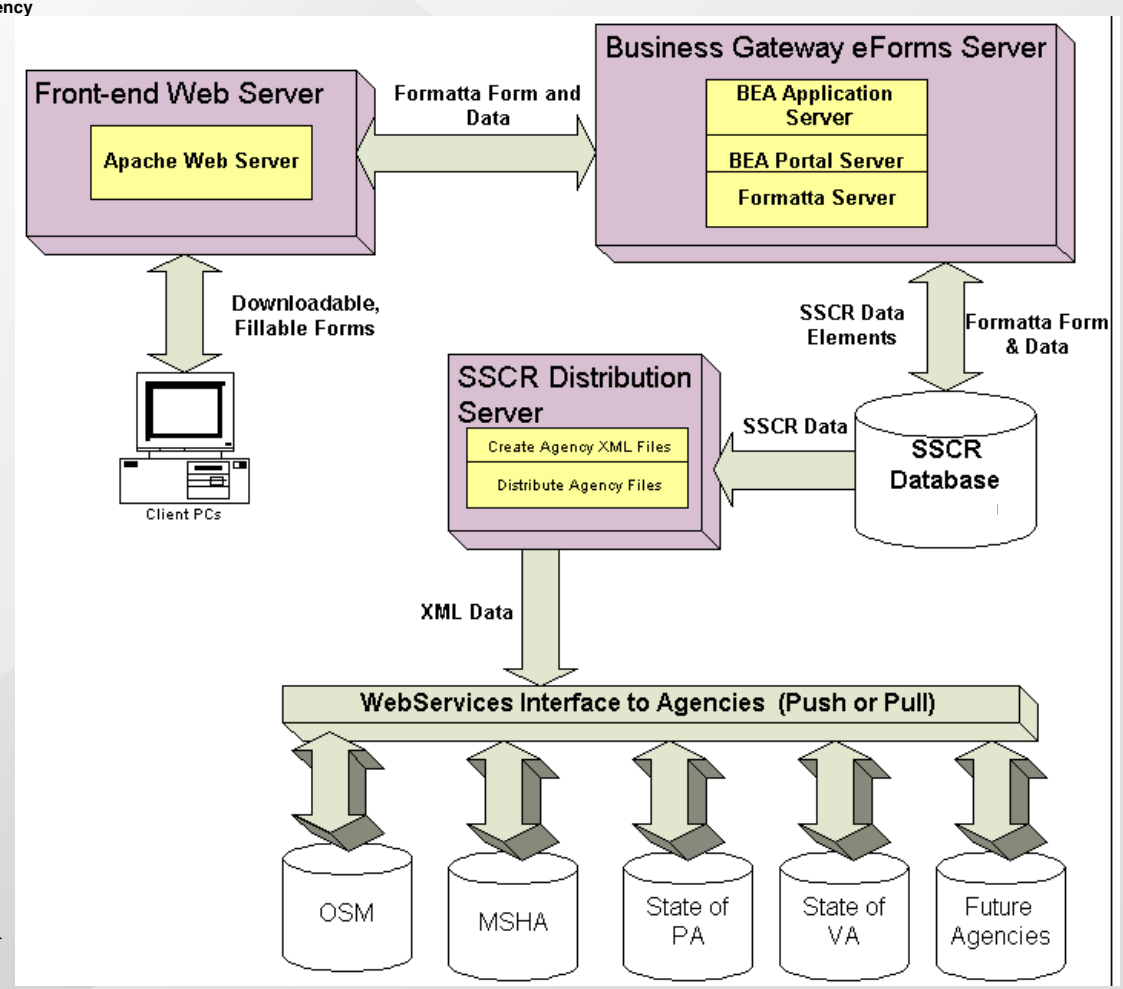
Imagine it. Done.

Agenda

- **SSCR Model Review**
  - Process Flow in e-Forms
- **Analysis**
  - Tasks 1-4
- **Recommendations**
- **Closing Summary Findings**

SSCR Model Review – Process Flow





UNISYS
Imagine it. Done.

Analysis – Task 1

- Evaluate suitability of the Business Gateway Surface Coal Mining Industry vertical's data standards for reuse throughout the entire Business Gateway initiative
  - The SSCR data standard does comply with OMB M-06-02 and the FEA DRM 2.0. Also, the SSCR data standards have been developed from ebXML core components, specifically:
    - ISO 15000-5 Business Information Entities and UN/CEFACT Core Components. This is consistent with both OMB policy (Circular A-119) and industry best practices.
  - In the Quarterly Mining Report Schema:
    - 65 element declarations are equivalent to ebXML CCTS BBIE
    - 34 element declarations are equivalent to ebXML CCTS ASBIE
    - 36 complex type element declarations are derived from ebXML CCTS ABIE
    - 53 complex type element declarations are derived from ebXML Unqualified Data Types

Analysis – Task 1

Strengths

- Via the core component structured design model of SSCR, the SSCR data model is compliant with the DRM V2.0, concepts, terms, and approaches for Data Sharing, Data Description, and Data Context
- Manifestation of structured design is realized through the Quarterly Mining Report XML exchange package structure
- Documented Data Dictionary
  - Coal Mining Core Component Analysis
  - Structure aligned with ISO 11179, contains:
    - Object Classes
    - Property Terms
    - Representation Terms

Analysis – Task 1

Weakness

- Completeness
  - 38% of data elements in Data Dictionary had omitted representation terms, definitions incomplete
- Documentation
  - Lack of conceptual and logical data models
    - UML
- Some use of DTD
  - DTDs do not support data types and namespaces
  - Should use XSD across board

Analysis – Task 2

- **Identify specific data elements within the Surface Coal Mining Industry vertical's data standards for inclusion in a common set of reusable data elements applicable to all of Business Gateway's verticals**
- **Challenge**
  - Highly specialized vertical industry (Coal), minimal translations to standard business vocabularies
- **Candidate Entities**
  - Organization: Name, Address, Identification, TaxIdentification
  - Address: LineOne, LineTwo, PostOfficeBox, CityName, PostCode, CountrySubDivision, Country
  - Country: Name
  - CountrySubDiv: Identification
  - Facility: Name, Address
  - Person: GivenName, FamilyName, MiddleName, Title, Communication
  - Communication: URL, CommunicationNumber
  - CommunicationNumber: Text
  - Contact: Identification, Person
  - Location: Directions
  - Report: Identification, Type, ReportCompletedDate, PeriodBeginningDate, PeriodEndingDate, YearDate, QuarterNumeric
  - Fee: Rate, Amount
  - Metrics: JobTypeCode
  - Metrics: TotalEmployeeQuantity
  - Metrics: TotalWagesAmount
  - PaymentMeans: Type

Analysis – Task 2

- **Strengths**
  - ebXML core component structure has been re-used
- **Weakness**
  - Customized SSCR design
  - XML Schema only imported UnqualifiedDataTypeSchemaModule

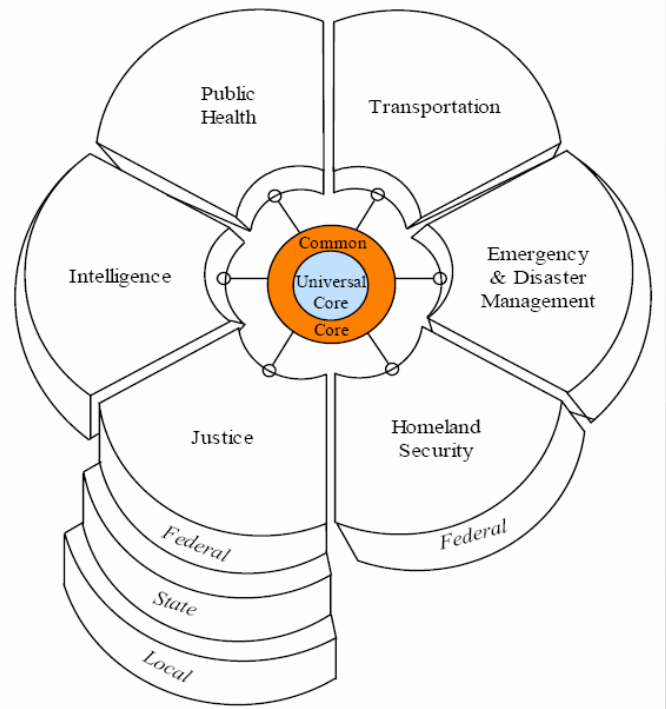
UNISYS
Imagine it. Done.

Analysis – Task 3

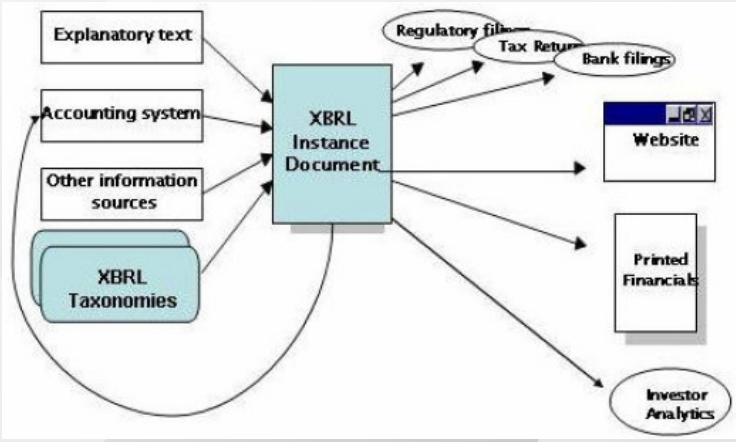
- Compare and contrast the preliminary set of "common data elements" extracted from the Business Gateway Surface Coal Mining Industry data standards with similar elements found in other VCS
- VCS Standards – Differences noted in all standards reviewed, data element comparison in access database to be provided which enumerates spread in entity and attribute semantics
- VCS Standards reviewed:
  - **NIEM**
  - **XBRL**
  - **UBL**
  - **ebXML**

Analysis – Task 3

- **Summary** - NIEM (http://www.niem.gov/)
  - Exchange package fundamentals for law enforcement domain, universal core components are applicable for SBA consideration
  - Primarily activity, event, location, person, document, organization, case entities, etc…
  - Not inherently business oriented vocabulary



NIEM Information Exchange Model



XBRL Instance Document Model

- **Summary** - XBRL (http://www.xbrl.org/Home/)
  - XBRL is in use for many different types of financial reporting – internal business reporting using general ledger-level accounting information, bank position reports consisting of thousands of individual data points collected by regulatory authorities, and financial reports consisting not only of individual financial numbers, but also textual disclosures of policies, tabular schedules of assets, consolidations, and a myriad of notes under a variety of accounting standards including US GAAP (Generally Accepted Accounting Principles)
  - XBRL captures business reporting information as an *XBRL Instance Document*. The information in an instance document is about a particular business entity, for some specific time periods, under a given set of definitions. The definitions appear in an *XBRL Taxonomy* such as the taxonomy for US GAAP financial statements, etc. The instance document contains information from many different sources, and information from it can be extracted to be presented and used in many different ways

Analysis – Task 3

- **Summary -** OASIS UBL (http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ubl)
- OASIS (Organization for the Advancement of Structured Information Standards) is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards
- The consortium produces more Web services standards than any other organization along with standards for security and e-business
- UBL 1.0 provides the following:
  - A library of XML schemas for reusable data components such as “Address,” “Item,” and “Payment” — the common data elements of everyday business documents
  - A set of XML schemas for common business documents such as “Order,” “Despatch Advice,” and “Invoice” that are constructed from the UBL library components and can be used in a generic order-to-invoice trading context
  - Support for the customization of UBL
  - To aid in deployment, standard UBL schemas are accompanied by a multitude of informative supporting materials as informative appendices. Materials include:
    - UML class diagrams of the document components on which the schemas are based
    - UML class diagrams describing all the document assemblies
    - Spreadsheet models defining the document assemblies
    - Formatting specifications for rendering all of the documents in the example use cases
- **Expanded UBL 2.0 soon to be released**

Analysis – Task 3

- **Summary -** ebXML (http://www.ebxml.org/)
  - ebXML (Electronic Business using eXtensible Markup Language), is a modular suite of specifications that enables enterprises of any size and in any geographical location to conduct business over the Internet. Using ebXML, companies now have a standard method to exchange business messages, conduct trading relationships, communicate data in common terms and define and register business processes
  - A joint 18-month effort, concluding in May 2001, of OASIS and UN/CEFACT (United Nations Centre for Trade Facilitation). Developed CCTS
  - The Core Components Technical Specification (CCTS) defines a syntax-neutral *metamodel* for business semantics
- ebXML Value
  - Provides the only globally developed open XML-based Standard built on a rich heritage of electronic business experience
  - Creates a Single Global Electronic Market Enables all parties irrespective of size to engage in Internet-based electronic business. Provides for plug and play shrink-wrapped solutions
  - Enables parties to complement and extend current EC/EDI investment and expand electronic business to new and existing trading partners
  - Facilitates convergence of current and emerging XML efforts
- **ebXML delivers the value by**
  - ***Developing technical specifications for the open ebXML infrastructure***
  - Creating the technical specifications with the world's best experts
  - Collaborating with other initiatives and standards development organizations

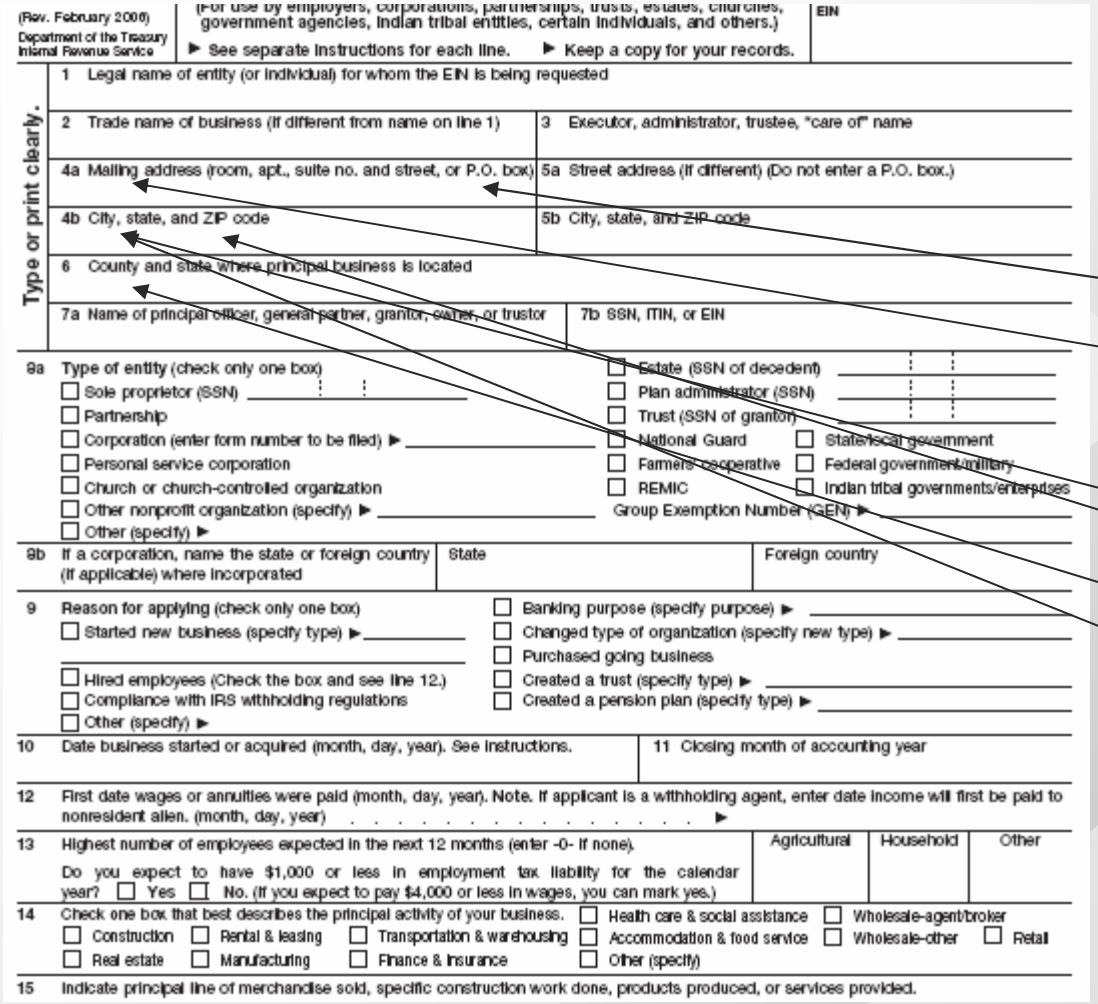
Analysis – Task 3

	Pros	*Cons*
NIEM	> Government Effort > Exchange package structure	> LE focused > Lacking ISO 11179 DD
XBRL	> Wide adoption in finance and accounting space > FDIC and FCC utilization	> Semantics optimized for finance and accounting > Lacking ISO 11179 DD
OASIS UBL	> Superset of ebXML core components > Robust documentation > Supporting models > Continued business expansion > Support by consortium that drives development, convergence, and adoption of e-business standards	> Limited V1 package structure (10)
ebXML	> Developed and maintains CCTS > UN/CEFACT support > Meta-model for business semantics defined > SSCR utilization	> Primarily core component dictionary support

Recommendations

- **Adopt UBL as standard for Business Gateway canonical data model**
  - As verticals conduct data harmonization, they should to the max extent possible harmonize on UBL standards and exchange schemas
  - Most suited to the semantics of business as compared to other VCS
  - Initial forms catalogue UBL matching metrics (UBL match of object class and property term union)
    - **51% exact match against forms catalogue name and description**
    - **84% partial match against forms catalogue name and description**
- **Develop vertical submission requirements and assessment strategy for Business Gateway funding**
  - Data dictionary development/maintenance based on ISO 11179 elements
  - UML conceptual and logical models for all UBL extensions
  - XML Schema submission for exchange packages
- **Perform additional analysis for required BG Universal Core**
  - What UBL data elements are used on every form
    - Gain access to index of individual form elements. Do per form matching
  - Provide verticals with starter set of mappings between OASIS UBL data elements and the BG forms catalogue to bootstrap harmonization
- **OASIS UBL appears to contain the most data elements relevant to the various Business Gateway domains and is available for royalty-free use**
  - Includes a complete ISO 11179-compliant data dictionary, various predefined, reusable exchange packages (encoded as W3C-compliant XML schemas), and UML diagrams of the overall UBL data model and its constituent component packages

Example – BG Universal Core Mapping



(Rev. February 2006)
Department of the Treasury
Internal Revenue Service

(For use by employers, corporations, partnerships, trusts, estates, churches, government agencies, Indian tribal entities, certain individuals, and others.)

▶ See separate instructions for each line. ▶ Keep a copy for your records.

EIN

Type or print clearly.

1 Legal name of entity (or individual) for whom the EIN is being requested

2 Trade name of business (if different from name on line 1)

3 Executor, administrator, trustee, "care of" name

4a Mailing address (room, apt., suite no. and street, or P.O. box)

5a Street address (if different) (Do not enter a P.O. box.)

4b City, state, and ZIP code

5b City, state, and ZIP code

6 County and state where principal business is located

7a Name of principal officer, general partner, grantor, owner, or trustor

7b SSN, ITIN, or EIN

9a Type of entity (check only one box)

- ☐ Sole proprietor (SSN)
- ☐ Partnership
- ☐ Corporation (enter form number to be filed) ▶
- ☐ Personal service corporation
- ☐ Church or church-controlled organization
- ☐ Other nonprofit organization (specify) ▶
- ☐ Other (specify) ▶
- ☐ Estate (SSN of decedent)
- ☐ Plan administrator (SSN)
- ☐ Trust (SSN of grantor)
- ☐ National Guard ☐ State/local government
- ☐ Farmers' cooperative ☐ Federal government/military
- ☐ REMIC ☐ Indian tribal governments/enterprises

Group Exemption Number (GEN) ▶

9b If a corporation, name the state or foreign country (if applicable) where incorporated — State — Foreign country

9 Reason for applying (check only one box)

- ☐ Started new business (specify type) ▶
- ☐ Hired employees (Check the box and see line 12.)
- ☐ Compliance with IRS withholding regulations
- ☐ Other (specify) ▶
- ☐ Banking purpose (specify purpose) ▶
- ☐ Changed type of organization (specify new type) ▶
- ☐ Purchased going business
- ☐ Created a trust (specify type) ▶
- ☐ Created a pension plan (specify type) ▶

10 Date business started or acquired (month, day, year). See instructions.

11 Closing month of accounting year

12 First date wages or annuities were paid (month, day, year). Note. If applicant is a withholding agent, enter date income will first be paid to nonresident alien. (month, day, year) ▶

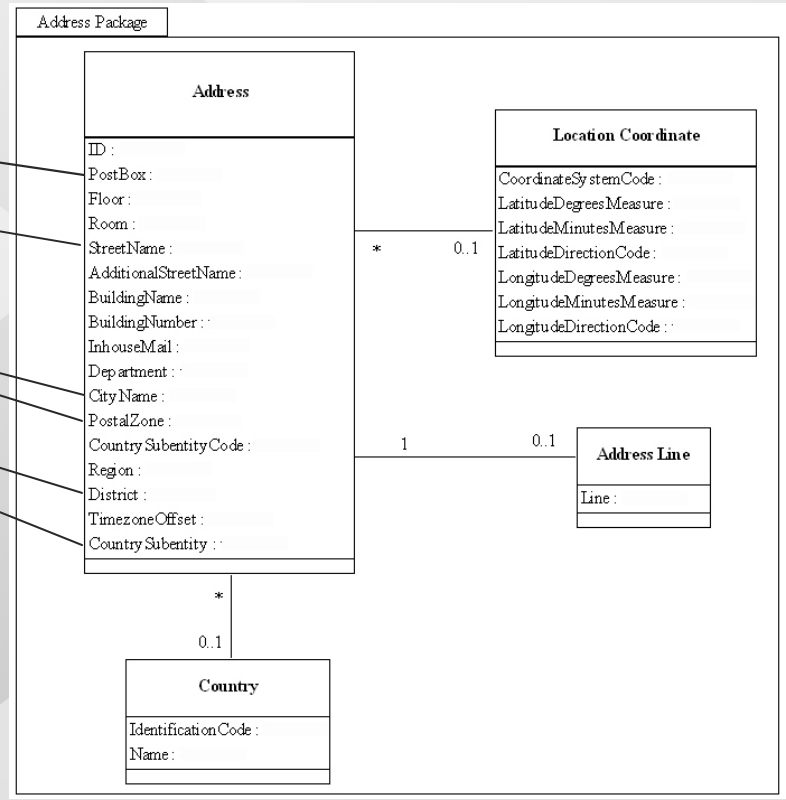
13 Highest number of employees expected in the next 12 months (enter -0- if none). — Agricultural — Household — Other

Do you expect to have $1,000 or less in employment tax liability for the calendar year? ☐ Yes ☐ No. (If you expect to pay $4,000 or less in wages, you can mark yes.)

14 Check one box that best describes the principal activity of your business.

- ☐ Construction ☐ Rental & leasing ☐ Transportation & warehousing ☐ Health care & social assistance ☐ Accommodation & food service ☐ Wholesale-agent/broker ☐ Wholesale-other ☐ Retail
- ☐ Real estate ☐ Manufacturing ☐ Finance & insurance ☐ Other (specify)

15 Indicate principal line of merchandise sold, specific construction work done, products produced, or services provided.



UNISYS
Imagine it. Done.

Summary Findings

- **SSCR did a good job of data harmonization**
  - Used UN/CEFACT ebXML appropriately (VCS)
- **Other Business Gateway verticals should follow the same harmonization process concepts**
  - Improvement – use OASIS UBL standard
    - OASIS UBL is a superset of ebXML, evolved and improved
    - More rigor on modeling and documentation
- **Business Gateway should issue Data Standards Policy Memo and consider recommendations to move the program forward**
- **Questions?**